Sample Design Options for the Survey of Apartments Participating in the World Trade Center Residential Cleanup Program

Authors:

Graham Kalton
David Ferraro
Lou Rizzo

Draft
March 19, 2004

Prepared for:

Environmental Protection Agency

Prepared by:

WESTAT
1650 Research Boulevard
Rockville, Maryland 20850

1. Introduction

The collapse of the Twin Towers of the World Trade Center (WTC) caused the incursion of contaminants into the indoor environment of buildings near to the site. In response, the Environmental Protection Agency (EPA) offered a clean-up program to approximately 23,000 apartments in an area of about one mile square in Manhattan. A total of 4,167 apartments in the clean-up area were cleaned under this voluntary program. A survey of a sample of these 4,167 apartments is now to be conducted to test for possible recontamination. The proposed focus of the survey is on possible asbestos recontamination. This report presents some sample design options for this survey.

The report is organized as follows. The next section provides some background information on the 4,167 apartments that participated in the WTO Residential Cleanup Program. Section 3 describes the objectives that the survey is designed to satisfy. Section 4 describes 16 possible sample designs for use in the survey. Section 5 discusses the problem that not all randomly selected apartments will provide data and describes the procedures to be applied when a sampled apartment fails to participate in the survey. Three appendices provide further details on the sample design.

2. Background on the 4,167 Apartments

Some details about the testing and cleaning of the 4,167 apartments in the Residential Cleanup Program are presented in Table 1. As can be seen from that table, most (3,893) of the apartments were tested by a modified aggressive sampling method, with the remaining 274 apartments being tested by an aggressive method. Most (3,386) were cleaned first and then tested, with the remainder (781) being tested first and cleaned only if necessary.

Table 1. Testing procedures for the 4,167 apartments in the cleanup program

	Sampling method		
Testing method	Modified aggressive	Aggressive	Total
Clean and test	3,146	240	3,386
Test	747	34	781
Total	3,893	274	4,167

Table 2 presents the outcome of the testing for the 4,167 apartments. Apartments were classified as in exceedance if their test results showed a level greater than the asbestos health benchmark of 0.0009f/cc (as measured by TEM)[1].

Table 2. Test results in the 4,167 apartments at initial testing

Test result	Sampling method		Total
	Modified aggressive	Aggressive	
Non detect	3,547	232	3,779
Detect	233	20	253
Exceedance	28	16	44
Overload	85	6	91
Total	3,893	274	4,167

Table 2 shows an overall exceedance rate of 1.1 percent among the 4,076 apartments for which asbestos levels could be determined. The exceedance rate is much higher, 5.9 percent, among the apartments subject to aggressive testing, than among apartments subject to modified aggressive testing, 0.7 percent. Although this may reflect a difference in the types of apartment undergoing the different kinds of cleaning, it is likely to be mainly the result of the different testing methods.

As noted earlier, most apartments were cleaned before testing, and others were tested first. Table 3 presents the test results for these two protocols with the modified aggressive testing procedure. The rate of exceedance among the apartments tested first (1.1%) is slightly, but not significantly, higher at 1.1 percent than the rate among apartments cleaned first (0.6%).

[1] This level corresponds to two or more long fibers on the filters. They were classified as 'detects' if there was one long fiber detected on the filters, and 'nondetects' if there were no long filters detected on the filters. Asbestos levels could not be determined for 91 apartments because of filter overload or other sampling problems. The residents of the apartments classified as exceedances and overloads were offered the opportunity to have their apartments cleaned or recleaned. Most accepted, and nearly all these apartments were cleaned on retestingNote that the classification used in Table 2 is slightly different from the procedure used to determine whether or not apartments were cleared during the dust Cleanup Program. The classification procedure used here ignores the filter samples that were not analyzed because of analytical or sampling errors other than overloads. The ignored samples include samples not analyzed for reasons such as lost samples, failed or damaged filters, and failure to meet QA standards for samples. Since there are no results for such samples—they provide neither qualitative nor quantitative information—they are not used in the classification. The overloads are included because they provide a qualitative result; dust levels in the apartments with an overload was relatively high.

Table 3. Test results for the two testing protocols using modified aggressive testing

Test result	Testing method		Total
	Clean and test	Test	
Non detect	2,870	677	3,547
Detect	194	39	233
Exceedance	20	8	28
Overload	62	23	85
Total	3,146	747	3,893

3. Survey objectives

Two primary objectives have been identified for the survey:

- To estimate the current asbestos exceedance rate for the 4,167 apartments participating in the World Trade Center Residential Cleanup Program; and
- To estimate the current asbestos exceedance rate for the subset of the apartments that are at risk of possible recontamination from shared air systems.

The sample size and sample design need to be able to produce estimates of adequate precision for each of these two domains. The second domain comprises the subset of apartments that have central HVAC systems and also apartments that have other kinds of shared air with other areas, classified as 'partial - common or exhaust'. It may have been of interest to produce separate estimates for apartments with central HVAC systems and those with other forms of shared air. However, only 472 apartments in the Cleanup Program have central HVAC systems, and 430 of them are in a single building (45 Wall Street); the other 42 are in only 9 other buildings. The small number of apartments with central HVAC systems, combined with their concentration in a few buildings, led to the decision that these apartments should not be treated as a separate domain for the survey.

Table 4 presents the numbers of apartments with central HVAC systems, with partial shared air, and with no shared air. The table also includes 1,275 apartments for which the shared air status is unknown. Most of these are likely to not have shared air.

Table 4. Air sharing for the 4,167 apartments

Air sharing	Number
Central HVAC	472
Partial-common or exhaust	2,396
No shared air	24
Unknown	1,275
Total	4,167

Two main forms of analysis are planned for each of the two domain estimates:

- A comparison of the domain exceedance rate with a background value for the exceedance rate for a ‘similar’ inner city area; and

- A comparison of the domain exceedance rate with the rate obtained from the original testing. For this form of analysis, the comparison will be made only for the subset of apartments that were originally tested by the modified aggressive method. Results for apartments tested by the aggressive method originally are not comparable with those that will be obtained from the modified aggressive method that will be used in the survey. It should also be noted that most of the apartments were cleaned before testing in the Cleanup Program, whereas they will be tested without cleaning in the survey.

In addition to the primary objectives listed above, data will be recorded on the number of long and short fibers detected on the filters of sampled apartments. These numbers will be used for additional analyses, including comparisons with the equivalent numbers recorded in the Cleanup Program.

4. Sample Design Options

The sample design for the survey is basically a simple one. A single-stage stratified sample of apartments will be drawn from the population of 4,167 apartments in the Cleanup Program. The two key decisions to be made are the overall sample size and the sample size allocated to the domain of apartments with shared air. These decisions are determined by the levels of precision required for the exceedance rates for the two domains. Some illustrative calculations are presented below to help guide these decisions.

First, consider the stratification factors that will be employed. These stratification factors are derived from the data collected in the initial testing. One group of particular interest is the 44 apartments that classified as in exceedance in the Cleanup Program. It is proposed that all these apartments be included in the survey, irrespective of the overall sample size for the survey. Another such group is the 91 apartments that were classified as overload in the Cleanup Program. It is proposed that they also all be included in the survey if the overall sample size is fairly large (say, 750 or more). It should be noted that it will not prove possible to collect the survey data from all the apartments in these two groups. For planning purposes, the assumption has been made that there will be a response rate of 60 percent.

Once the above groups have been removed, the primary stratification for the remainder is by shared air status as given in Table 4. This stratification provides the means for controlling the sample size of the shared air domain; different sampling rates can be used in different primary strata to yield required sample sizes. Within the four primary strata, further stratification will be incorporated based on the original testing protocol and the detect/nondetect results. Finally, the list of apartments in each ultimate stratum will be placed in a relevant order (e.g., by geography) so that the application of systematic sampling will yield an 'implicit' stratification according to the ordering.

A constant sampling fraction will be used within each of the substrata within a primary stratum. The use of a constant sampling fraction yields what is termed proportionate stratification. Proportionate stratification has the attractive feature that it can only improve the precision of survey estimates as compared with an unstratified design. It cannot cause a loss of precision.

Without information about current exceedance rates for the different strata, a proportionate allocation of the sample across the primary strata would also be appropriate for estimating the overall exceedance rate under the first objective listed earlier. However, a proportionate allocation may not yield a large enough sample of apartments with shared air systems to produce an estimate of the exceedance rate of adequate precision for the second domain. Thus a disproportionate allocation, 'oversampling' apartments with shared air systems, may be required. Note that, for a fixed sample size, this oversampling of the second domain will result in less precision for the overall estimate than would have been achieved with a proportionate allocation.

As an aid to making decisions about the total sample size, and what proportion of the total sample size to allocate to the shared air domain, we examine four possible sample sizes (250, 500, 750 and 1,000) and four rates of oversampling for the shared air domain (no oversampling—i.e., a proportionate allocation and an oversampling ratio of 1.0—and sampling the shared air domain at 1.5, 2.0 and 2.5 times the rate used for the apartments with no shared air). The sample sizes that result from the 16 combinations of these two factors are displayed in Table 5. As the table shows, for a given overall sample size, an increase in the oversampling rate for the shared air domain reduces the sample size for apartments with no shared air. As a result, the estimate of the exceedance rate for the apartments with no shared air becomes less precise, which in turn lowers the precision of the estimate of the exceedance rate for the total population of apartments in the Cleanup Program. The tables in Appendix A present the sample sizes for each of the substrata for the 16 sample design options

Table 5. Sample sizes for various combinations of total sample size and the oversampling rate for the shared air domain (S) as compared with the rate for apartments with no shared air (NS)

Oversampling rate for shared air domain	Sample size							
	250		500		750		1,000	
	S	NS	S	NS	S	NS	S	NS
1.0	170	80	339	161	509	241	679	321
1.5	190	60	380	120	570	180	760	240
2.0	202	48	404	96	606	144	809	191
2.5	210	40	420	80	631	119	841	159

To illustrate the magnitude of the sampling errors resulting from the above 16 sample designs, we have computed approximate standard errors for the sample estimate of an exceedance rate of 1 percent (see Appendix B for details). These computations assume that the proportionate stratification does not reduce the standard error to an appreciable extent. The resultant standard errors may therefore be overestimates, but the extent of overestimation is likely to be negligible. Table 6 presents the standard errors for the total domain estimate (T) and the shared air domain (S) for each of the 16 sample designs. Note that with an oversampling rate of 2.5, the standard errors for the estimated exceedance rates for the total domain and for the shared air domain are about the same.

Table 6. Standard errors for an estimate of an exceedance rate of 1 percent for the total population (T) and the shared air domain (S) for the 16 sample designs

Oversampling rate for shared air domain	Oversampling rate							
	250		500		750		1,000	
	T	S	T	S	T	S	T	S
1.0	0.61%	0.74%	0.42%	0.51%	0.33%	0.40%	0.27%	0.33%
1.5	0.62	0.70	0.43	0.48	0.34	0.37	0.28	0.31
2.0	0.64	0.67	0.44	0.46	0.35	0.36	0.29	0.30
2.5	0.67	0.66	0.46	0.45	0.37	0.35	0.31	0.29

Table 6 shows how the sampling error decreases with increasing sample size. For a given total sample size, the sampling error of the estimate for the shared air domain decreases with higher rates of oversampling; however, the sampling error of the estimate for the total population increases with higher rates of oversampling. The sample design to be chosen should be the one that satisfies the precision requirements for both domain estimates with the smallest sample size.

The interpretation of the standard error involves considering the variation in the sample estimates that would occur if the same sampling procedure were repeated an infinite number of times. The standard error is then the square root of the average of the squared deviations of the sample estimates from the true population value. With large samples, approximately 95 percent of the sample values will fall in the interval P - 1.96SE to P + 1.96SE, where P is the population value and SE is the standard error. Although this approximation is not entirely adequate when both P and the sample size are small, it can nevertheless provide useful guidance.

Suppose, for example, that the true exceedance rate is 1.0 percent in a population of 4,200 (rounded for simplicity), i.e., 42 of the apartments are in exceedance. Consider a sample of 1,000 apartments with no oversampling (i.e., an oversampling rate of 1.0). Then, from Table 6 the standard error for the sample proportion in the total domain is SE = 0.27. Thus approximately 95 percent of the sample estimates from the repeated sampling would fall in the interval from 1.0 - (1.96 x 0.27) to 1.0 + (1.96 x 0.27) or from 0.47 to 1.53 percent. Alternatively expressed, 95 percent of the samples would have between 5 and 15 apartments in exceedance. Based on these numbers, 95 percent of the samples would estimate the number of 4,200 apartments in exceedance as between about 20 and 64.

Consider now the same design but with a sample of 500. The standard error from Table 6 is now 0.42. In this case, the interval within which 95 percent of the sample estimates would fall is widened to 0.18 to 1.82 percent and the corresponding interval for the estimated number of the 4,200 apartments in exceedance would be from 7 to 77.

It should be noted that the magnitude of the standard error is highly dependent on the true exceedance rate. Table 7 illustrates this point by providing the standard errors corresponding to those in Table 6 for exceedance rates of 0.5 percent, 2.0 percent and 3.0 percent. The standard errors for a given sample size are approximately $\sqrt{2} = 1.41$ times larger for an exceedance rate of 2.0 percent and approximately $\sqrt{2}$ times smaller for an exceedance rate of 0.5 percent than for one of 1.0 percent. They are approximately $\sqrt{3} = 1.73$ times larger for an exceedance rate of 3.0 percent than for one of 1.0 percent.

Table 7. Standard errors for an estimate of exceedance rates of 0.5 percent, 2 percent and 3 percent for the total populations (T) and the shared air domain (S) for the 16 sample designs

Oversampling rate for shared air domain	Oversampling rate							
	250		500		750		1,000	
	T	S	T	S	T	S	T	S
Population exceedance rate of 0.5 percent								
1.0	0.43%	0.53%	0.30%	0.36%	0.23%	0.28%	0.19%	0.24%
1.5	0.44	0.49	0.30	0.34	0.24	0.26	0.20	0.22
2.0	0.46	0.48	0.31	0.33	0.25	0.25	0.21	0.21
2.5	0.48	0.47	0.33	0.32	0.26	0.25	0.22	0.20
Population exceedance rate of 2.0 percent								
1.0	0.86%	1.04%	0.59%	0.71%	0.46%	0.56%	0.39%	0.47%
1.5	0.88	0.98	0.60	0.67	0.47	0.52	0.40	0.43
2.0	0.91	0.95	0.62	0.65	0.49	0.50	0.41	0.42
2.5	0.94	0.93	0.65	0.63	0.52	0.49	0.43	0.41
Population exceedance rate of 3.0 percent								
1.0	1.05%	1.27%	0.72%	0.87%	0.56%	0.69%	0.47%	0.57%
1.5	1.07	1.20	0.73	0.81	0.58	0.64	0.48	0.53
2.0	1.11	1.16	0.76	0.79	0.60	0.61	0.50	0.51
2.5	1.15	1.13	0.79	0.77	0.63	0.60	0.53	0.49

In practice, of course, only one sample is selected, and the issue becomes one of placing a range around the sample estimate within which the true value lies, termed a confidence interval. Computation of confidence intervals in the case of small proportions requires special techniques. Appendix B provides a method for computing approximate confidence intervals for this case. Table 8 provides an excerpt from the table in that Appendix. The table is arranged in blocks with estimates as close to 0.5 percent, 1.0 percent, 2.0 percent and 3.0 percent as possible with the different sample sizes.

Table 8. Illustrations of approximate confidence intervals for given sample sizes and numbers of sampled apartments in exceedance

Sample size	Number in exceedance	Estimate	95 percent confidence interval	
			Lower bound	Upper bound
Estimate of about 0.5%				
250	1	0.40%	0.01%	2.16%
500	2	0.40	0.06	1.37
750	4	0.53	0.18	1.27
1,000	5	0.50	0.20	1.08
Estimate of about 1.0%				
250	3	1.20%	0.27%	3.40%
500	5	1.00	0.34	2.24
750	8	1.07	0.51	2.00
1,000	10	1.00	0.53	1.74
Estimate of about 2.0%				
250	5	2.00%	0.68%	4.54%
500	10	2.00	1.01	3.54
750	15	2.00	1.20	3.17
1,000	20	2.00	1.32	2.94
Estimate of about 3.0%				
250	7	2.80%	1.18%	5.61%
500	15	3.00	1.75	4.77
750	22	2.93	1.94	4.28
1,000	30	3.00	2.15	4.11

As an illustration, the first row of the table indicates that if one apartment out of a sample of 250 apartments (sampled without oversampling) are in exceedance, the estimate of the exceedance rate is 0.4 percent. The 95 percent confidence interval for the true exceedance rate from this small sample is a wide one, from 0.01 percent to 2.16 percent. The widths of the confidence intervals decrease as the sample size increases.

The above discussion has focused on the estimation of the exceedance rate and computing a confidence interval for it. Another common form of statistical analysis is to perform a significance test to determine whether the sample estimate is significantly different from another value, such as a background rate from some other source or the rate as determined from the Cleanup Program. Some results on the power of significance tests to detect true differences of specified magnitudes are given in Appendix C.

5. Nonresponse

The quality of the sample depends on the selection of a well-designed probability sample *and* the collection of the survey data from the selected sample. Nonresponse is a threat to the validity of the survey estimates. Strenuous efforts will therefore be made to minimize nonresponse, which will occur mainly from refusals to participate and not-at-homes. Nonresponse will also arise if the filters are obtained but cannot be analyzed. As an incentive, residents will be offered the opportunity to have their apartments cleaned if they are found to be in exceedance.

In order to maintain the intended sample size, substitute apartments will be surveyed as replacements for nonresponding apartments. Each substitute apartment will be carefully selected to be similar to its nonresponding apartment. Substitute apartments will be selected from the same substratum as the nonresponding apartments, and to the extent possible as the adjacent apartment in the ordered list in that substratum. This matching of substitute and nonresponding apartments aims to reduce the bias in the survey estimates that can arise from nonresponse. The use of substitution does not replace the need to make every effort to collect data from the original sample.

Substitution is not applicable in the case of the groups in which all the apartments are selected for the survey. For planning purposes, a response rate of 60 percent has been assumed for these groups. The sample losses from nonresponse in these groups will be allocated to the rest of the sample in order to maintain sample size. Nonresponse weighting adjustments will be used in the analyses in order that these groups are not underrepresented.

Appendix A

EXPECTED SAMPLE SIZE ALLOCATIONS UNDER THE ALTERNATIVE SAMPLE DESIGNS

The tables below present the sample size allocations expected for the various implicit strata within the primary strata for the 16 sample designs (four oversampling rates 1.0, 1.5, 2.0, 2.5; four sample sizes 250, 500, 750, 1,000). The oversampling rate is the rate at which the shared air stratum is oversampled as compared to the complement stratum. Table A-1 presents the expected sample size allocations for no oversampling, and Tables A-2 through A-4 present the expected sample size allocations for oversampling rates 1.5, 2.0, and 2.5 respectively. All of these designs include selecting all of the apartments in the exceedance substrata for all sample sizes (with an assumed 60% response rate), and also selecting all the apartments in the overload substrata for the larger sample sizes of 750 and 1,000 (with an assumed 60% response rate).

Table A-1. Sample size allocation to sampling strata under proportional allocation to the primary strata, and using four overall sample sizes

Primary stratum	Clean vs. Test	Apt Classification	Baseline population size	Sample size 250	Sample size 500	Sample size 750	Sample size 1,000
Shared air	Clean and test	NonDetect + Detect	2,344	127	269	389	534
		Overload	42	2	5	25	25
		Exceedance	25	15	15	15	15
	Test only	NonDetect + Detect	444	24	51	74	101
		Overload	10	1	1	6	6
		Exceedance	3	2	2	2	2
		Total	2,868	171	343	511	683
Complement	Clean and test	NonDetect + Detect	941	51	108	156	214
		Overload	24	1	3	14	14
		Exceedance	10	6	6	6	6
	Test only	NonDetect + Detect	303	16	35	50	69
		Overload	15	1	2	9	9
		Exceedance	6	4	4	4	4
		Total	1,299	79	157	239	317
Both Strata	Both strata	Grand total	4,167	250	500	750	1000
		NonDetect + Detect	4,032	219	463	669	919
		Overload	91	5	10	55	55
		Exceedance	44	26	26	26	26

Table A-2. Sample size allocation to sampling strata under the 1.5 times oversampling plan, and using four overall sample sizes

Primary stratum	Clean vs. Test	Apt Classification	Baseline population size	Sample size 250	Sample size 500	Sample size 750	Sample size 1,000
Shared air	Clean and test	NonDetect + Detect	2,344	141	299	434	596
		Overload	42	3	5	25	25
		Exceedance	25	15	15	15	15
	Test only	NonDetect + Detect	444	27	57	82	113
		Overload	10	1	1	6	6
		Exceedance	3	2	2	2	2
		Total	2,868	188	380	564	756
Complement	Clean and test	NonDetect + Detect	941	38	80	116	159
		Overload	24	1	3	14	14
		Exceedance	10	6	6	6	6
	Test only	NonDetect + Detect	303	12	26	37	51
		Overload	15	1	2	9	9
		Exceedance	6	4	4	4	4
		Total	1,299	62	120	186	244
Both Strata	Both strata	Grand total	4,167	250	500	750	1,000
		NonDetect + Detect	4,032	218	462	669	919
		Overload	91	5	12	55	55
		Exceedance	44	26	26	26	26

Table A-3. Sample size allocation to sampling strata under the 2.0 times oversampling plan, and using four overall sample sizes

Primary stratum	Clean vs. Test	Apt Classification	Baseline population size	Sample size 250	Sample size 500	Sample size 750	Sample size 1,000
Shared air	Clean and test	NonDetect + Detect	2,344	150	317	460	632
		Overload	42	3	6	25	25
		Exceedance	25	15	15	15	15
	Test only	NonDetect + Detect	444	28	60	87	120
		Overload	10	1	1	6	6
		Exceedance	3	2	2	2	2
		Total	2,868	198	401	595	799
Complement	Clean and test	NonDetect + Detect	941	30	64	92	127
		Overload	24	2	3	14	14
		Exceedance	10	6	6	6	6
	Test only	NonDetect + Detect	303	10	20	30	41
		Overload	15	1	2	9	9
		Exceedance	6	4	4	4	4
		Total	1,299	52	99	155	201
Both Strata	Both strata	Grand total	4,167	250	500	750	1,000
		NonDetect + Detect	4,032	218	461	669	919
		Overload	91	6	12	55	55
		Exceedance	44	26	26	26	26

Table A-4. Sample size allocation to sampling strata under the 2.5 times oversampling plan, and using four overall sample sizes

Primary stratum	Clean vs. Test	Apt Classification	Baseline population size	Sample size 250	Sample size 500	Sample size 750	Sample size 1,000
Shared air	Clean and test	NonDetect + Detect	2,344	155	329	477	656
		Overload	42	3	6	25	25
		Exceedance	25	15	15	15	15
	Test only	NonDetect + Detect	444	29	62	90	124
		Overload	10	1	1	6	6
		Exceedance	3	2	2	2	2
		Total	2,868	205	415	616	828
Complement	Clean and test	NonDetect + Detect	941	25	53	77	105
		Overload	24	2	3	14	14
		Exceedance	10	6	6	6	6
	Test only	NonDetect + Detect	303	8	17	25	34
		Overload	15	1	2	9	9
		Exceedance	6	4	4	4	4
		Total	1,299	45	85	134	172
Both Strata	Both strata	Grand total	4,167	250	500	750	1,000
		NonDetect + Detect	4,032	218	461	669	919
		Overload	91	6	13	55	55
		Exceedance	44	26	26	26	26

Appendix B

SAMPLING ERROR CALCULATIONS

The standard error of a proportion (*p*) for a simple random sample (SRS) of size *n* sampled without replacement from a population of size *N* is given by

$$SE(p) = \sqrt{\frac{P(1-P)}{n} \cdot \frac{N-n}{N-1}}$$

where *P* is the population proportion with the specified attribute (Kish, 1965, equation [2.4.6]). The sampling distribution on which this formula is based is the hypergeometric distribution. Although the proposed sample design is a stratified sample design rather than a simple random sample, the SRS standard error formula can serve as a useful approximation.

In some of the alternative designs, the sample is proportionately allocated across strata, while in others it is disproportionately allocated between the shared air domain and the remainder. With proportionate stratification, the standard error will be smaller than that given by the SRS formula above. However, the gains in precision from proportionate stratification are typically modest, so that using the SRS formula will only slightly overestimate the true variance. Thus, the SRS formula has been used for cases of proportionate allocation. These cases include all cases of estimates for the shared air domain and the case with the oversampling rate of 1.0 for the total domain. (Strictly, allowance should be made for the groups of exceedances and overloads sampled with certainty, but they are so small that this allowance would have a negligible effect.)

For estimates for the total domain where oversampling was applied, the sample is a disproportionate one. Under assumptions of homogeneity of strata with respect to exceedance rates, the disproportionate allocation leads to a loss of precision. The standard error is increased over the SRS standard error by a factor of

$$\sqrt{1 + W(1-W)\frac{K-1}{K}}$$

where *K* is the rate of oversampling and *W* is the proportion of the total population in the oversampled domain (Kish, 1965, equation [11.7.8]). This factor has been incorporated into the standard error computations where needed. The results of these computations are presented in Tables 6 and 7 in the main text.

The standard error results in Tables 6 and 7 are useful for planning a sample design, but cannot be used directly in assessing the precision of a sample estimate since the true population is not known. In order to provide error bounds for the sample estimate within which the population value will lie with high probability (conventionally 95 percent probability), we have computed approximate 95 percent confidence intervals using the hypergeometric distribution (with adjustments for the "effective sample size" in the case of oversampling). The results are intended to indicate only approximate levels of precision. The technique used to construct confidence intervals in analyzing the survey data will incorporate allowances for the stratification and weighting for unequal selection probabilities and nonresponse.

Using the hypergeometric distribution, the probabilities of *k* apartments in exceedance in a sample of size *n* is given by

$$P(k) = \frac{\binom{E}{k}\binom{\overline{E}}{n-k}}{\binom{N}{n}}$$

where E is the number of apartments in exceedance, and $\overline{E}$ is the number of apartments not in exceedance in the population, and $N = E + \overline{E}$ is the total number of apartments in the population (Cochran, 1977, Section 3.5). Table B-1 illustrates the hypergeometric distribution for the case of N = 4,200 apartments with E = 42 in exceedance and $\overline{E} = 4{,}158$ not in exceedance, i.e., an exceedance rate of 1.0 percent, and for the four sample sizes (n) under consideration.

Table B-1. Hypergeometric sampling distributions for various sample sizes from a population of 4,200 apartments with 1.0 percent in exceedance

Number of exceedances in the sample	Sample size							
	250		500		750		1,000	
	Sample estimate	Probability	Sample estimate	Probability	Sample estimate	Probability	Sample estimate	Probability
0	0.0%	7.5%	0.0%	0.5%	0.0%	0.0%	0.0%	0.0%
1	0.4	20.1	0.2	2.7	0.1	0.2	0.1	0.0
2	0.8	26.3	0.4	7.6	0.3	1.0	0.2	0.1
3	1.2	22.2	0.6	13.8	0.4	3.0	0.3	0.4
4	1.6	13.7	0.8	18.3	0.5	6.4	0.4	1.1
5	2.0	6.5	1.0	18.8	0.7	10.6	0.5	2.7
6	2.4	2.5	1.2	15.6	0.8	14.3	0.6	5.3
7	2.8	0.8	1.4	10.8	0.9	16.1	0.7	8.6
8	3.2	0.2	1.6	6.4	1.1	15.3	0.8	11.8
9	3.6	0.1	1.8	3.2	1.2	12.5	0.9	13.9
10	4.0	0.0	2.0	1.4	1.3	9.0	1.0	14.4
11	4.4	0.0	2.2	0.6	1.5	5.6	1.1	13.1
12	4.8	0.0	2.4	0.2	1.6	3.2	1.2	10.5
13	5.2	0.0	2.6	0.1	1.7	1.6	1.3	7.6
14	5.6	0.0	2.8	0.0	1.9	0.7	1.4	4.9
15	6.0	0.0	3.0	0.0	2.0	0.3	1.5	2.8
16+	--	0.0	--	0.0	--	0.0	--	2.7
Total		100.0		100.0		100.0		100.0

As can be seen from Table B-1, with a sample of 250, 96 percent of the sample estimates of the true rate of 1.0 percent fall within the range from 0.0 percent to 2.0 percent. Note that there is a 7.5 percent probability of a sample estimate of no exceedances with this sample size. The range of likely sample estimates decreases with sample size. With a sample of 1,000, 96 percent of the sample estimates fall within the range from 0.5 percent to 1.5 percent.

An exact 95 percent confidence interval procedure takes as its input the population size *N*, the sample size *n*, and the exceedance rate, $p = k/n$. Then, we find all values *P* such that the realized incidence outcome *p* lies between the 2.5th and 97.5th percentiles of the distribution of *p* with *P* as the true incidence rate (Cochran, 1977, Section 3.6).

Population size = *N*; Sample size = *n*; *k*~Hypergeometric (*N, n, P*).
Sample incidence result $p = k/n$.
We set 95% confidence interval $[P_0, P_1]$ to contain all values P' such that $q_{.025}(P') \le p \le q_{.975}(P')$, where $q_{.025}(P')$ is the 0.025 percentile of the sample incidence *p* under the Hypergeometric (*N, n,* P'), and $q_{.975}(P')$ is the 0.975 percentile of the sample incidence *p* under the Hypergeometric (*N, n,* P').

Table B-2 presents 95 percent confidence intervals for a selection of sample outcomes for the case of a population of size $N = 4{,}200$ and for the various sample sizes. The chosen sample outcomes are the likely outcomes for the case of a population exceedance rate of 1.0 percent. However, the confidence intervals do not involve any assumptions about the population exceedance rate.

References

Cochran, W. G. (1977). *Sampling Techniques*. 3rd ed. New York:Wiley.

Kish, L. (1965). *Survey Sampling*. New York:Wiley.

Table B-2. Lower and upper bounds for 95 percent confidence intervals based on the hypergeometric distribution for various sample outcomes and sample sizes

Sample size	Number of exceedances in the sample	Sample estimates	95 percent confidence interval	
			Lower bound	Upper bound
250	0	0.00%	0.00%	1.42%
	1	0.40	0.01	2.16
	2	0.80	0.11	2.81
	3	1.20	0.27	3.40
	4	1.60	0.46	3.97
	5	2.00	0.68	4.54
	6	2.40	0.92	5.06
500	1	0.20	0.01	1.06
	2	0.40	0.06	1.37
	3	0.60	0.13	1.69
	4	0.80	0.25	1.98
	5	1.00	0.34	2.24
	6	1.20	0.49	2.53
	7	1.40	0.61	2.78
	8	1.60	0.72	3.04
	9	1.80	0.87	3.31
	10	2.00	1.0	3.54
750	2	0.27	0.04	0.89
	3	0.40	0.11	1.08
	4	0.53	0.18	1.27
	5	0.67	0.25	1.47
	6	0.80	0.32	1.66
	7	0.93	0.42	1.82
	8	1.07	0.51	2.00
	9	1.20	0.61	2.18
	10	1.33	0.70	2.36
	11	1.47	0.80	2.52
	12	1.60	0.89	2.68
	13	1.73	0.99	2.86
1,000	4	0.40	0.13	0.94
	5	0.50	0.20	1.08
	6	0.60	0.25	1.20
	7	0.70	0.32	1.35
	8	0.80	0.39	1.47
	9	0.90	0.46	1.61
	10	1.00	0.53	1.74
	11	1.10	0.61	1.85
	12	1.20	0.68	1.97
	13	1.30	0.75	2.11
	14	1.40	0.84	2.23

Appendix C

POWER LEVELS OF TESTS FOR SIGNIFICANT DIFFERENCES

One form of analysis of the data from the survey will be to compare the sample exceedance rate with the exceedance rate from the Cleanup Program or with a background exceedance rate from other sources. This form of analysis requires a statistical significance test to determine whether the difference between the two rates could simply be due to sampling error.

To illustrate this kind of analysis, we consider a one-sided (directional) significance test that tests whether the survey exceedance rate is greater than a background rate. The null hypothesis is that the two rates are the same and the alternative hypothesis is that the true exceedance rate at the time of the survey is greater than the background rate. We assume that the background rate is obtained from an independent source and we use a conventional 5.0 percent significance level. In designing studies to test hypotheses, a power of 80 percent is commonly specified. The power represents the probability of correctly concluding that there is a real difference between the two population rates for a given difference in these rates.

Table C-1 presents the smallest population exceedance rate for a population of 4,200 apartments at the time of the survey to have a power of 80 percent for detecting that this rate differs from a true background rate of 1.0 percent. These power calculations are approximate ones based on large-sample normal distribution approximations and ignoring stratification and weighting. The sample design considered has a proportionate allocation across strata. The standard errors for the survey estimates are obtained as described in Appendix B. For illustrative purposes the background estimate is assumed to have a standard error of 0.5 percent. The standard error of the difference between the survey and background estimates is computed as the square root of the sum of the squared standard errors of the two estimated rates.

Table C-1. Smallest detectable incidence rates with 80 percent power using a one-sided test with a 5 percent significance level

Sample size	Standard error of survey exceedance rate	Standard error of the difference in the rates	Smallest exceedance rate with 80% power of being detected
250	0.61%	0.79%	1.97%
500	0.42	0.65	1.63
750	0.33	0.60	1.50
1,000	0.27	0.57	1.43
170	0.74	0.89	2.23
339	0.51	0.71	1.78
509	0.40	0.64	1.60
679	0.33	0.60	1.50

As can be seen in Table C-1, the sample size of 250 allows us to detect a significant difference between the background and the survey population rates with 80 percent power only if the survey population incidence is about two times as high as the background (1.97% vs. 1.0%). With a sample size of 1,000, we can detect a difference with 80 percent power when the survey population incidence is only 1.43 percent. Since the shared air domain has a smaller sample size, only larger differences can be detected with 80 percent power with this proportionate stratified design.